

R で統計解析入門

(15) 多重性の調整



準備：データ「DEP」の読み込み

1. データ「DEP」を以下からダウンロードする
<http://www.cwk.zaq.ne.jp/fkhud708/files/dep.csv>
2. ダウンロードした場所を把握する　ここでは「c:/temp」とする
3. R を起動し，2. の場所に移動し，データを読み込む

```
> setwd("c:/temp") # dep.csv がある場所に移動
> DEP <- read.csv("dep.csv") # dep.csv を読み込む
> DEP$GROUP <- factor(DEP$GROUP) # 薬剤の水準を 2 カテゴリに
> DEP$Y <- ifelse(DEP$EVENT==1, 1, 0) # あり 1,なし 0 なる変数を作成
> head(DEP)
```

	GROUP	QOL	EVENT	DAY	PREDRUG	DURATION	Y
1	A	15	1	50	NO	1	1
2	A	13	1	200	NO	3	1
3	A	11	1	250	NO	2	1
4	A	11	1	300	NO	4	1
5	A	10	1	350	NO	2	1
6	A	9	1	400	NO	2	1



準備：架空のデータ「DEP」の変数

- ▶ **GROUP**：薬剤の種類（A, B, C）
- ▶ **QOL**：QOL の点数（数値） 点数が大きい方が良い
- ▶ **EVENT**：改善の有無（1：改善あり，2：改善なし）
QOL の点数が 5 点以上の場合を「改善あり（イベント発生）」とする
- ▶ **Y**：改善の有無（1：イベント，0：打ち切り）
変数 **EVENT** の 2 を 0 に置き換えただけの変数
- ▶ **DAY**：観察期間（数値，単位は日）
- ▶ **PREDRUG**：前治療薬の有無（YES：他の治療薬を投与したことあり，
NO：投与したことなし）
- ▶ **DURATION**：罹病期間（数値，単位は年）



準備：架空のデータ「DEP」（一部）

GROUP	QOL	EVENT	DAY	PREDRUG	DURATION
A	15	1	50	NO	1
A	13	1	200	NO	3
A	11	1	250	NO	2
A	11	1	300	NO	4
A	10	1	350	NO	2
A	9	1	400	NO	2
A	8	1	450	NO	4
A	8	1	550	NO	2
A	6	1	600	NO	5
A	6	1	100	NO	7
A	4	2	250	NO	4
A	3	2	500	NO	6
A	3	2	750	NO	3
A	3	2	650	NO	7
A	1	2	1000	NO	8
A	6	1	150	YES	6
A	5	1	700	YES	5
A	4	2	800	YES	7
A	2	2	900	YES	12
A	2	2	950	YES	10
B	13	1	380	NO	9
B	12	1	880	NO	5
B	11	1	940	NO	2
B	4	2	20	NO	7
B	4	2	560	NO	2
B	5	1	320	YES	11
B	5	1	940	YES	3
B	4	2	80	YES	6
B	3	2	140	YES	7
B	3	2	160	YES	13



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. And の論理, Or の論理, ボンフェローニの方法
4. 固定順検定
5. ダネットの方法とテューキーの方法
6. その他の方法



【復習】 検定の手順

1. 比較の枠組みを決める
2. 比較するものの間に差がないという仮説（帰無仮説 H_0 ）を立てる
3. 帰無仮説とは裏返し（差がある）の仮説（対立仮説 H_1 ）を立てる
4. 帰無仮説が成り立つという条件の下で、手元にあるデータよりも極端なことが起こる確率（p 値）を計算する
5. 計算した確率が非常に小さい場合は「珍しいデータが得られた」と考えるのではなく「そんな珍しいことは通常起こらない・・・」
「帰無仮説 H_0 （差がないという仮説）自体が間違っている」と考え、対立仮説 H_1 が正しいと結論付ける
6. 計算した確率が小さくない場合は「帰無仮説 H_0 が間違っている」といえないので「帰無仮説 H_0 が間違っているとはいえない」と考える



検定を行う際の 2 種類の error

- ▶ **Type I error (α)** : 帰無仮説 H_0 は正しいが, 検定の結果「帰無仮説 H_0 は間違い」とする確率
- ▶ **Type II error (β)** : 帰無仮説 H_0 は間違いだが, 検定の結果「帰無仮説 H_0 は正しい」とする確率
- ▶ 上記に関連したものとして「有意水準」と「検出力」がある
 - ▶ **有意水準 (α)** : 検定の結果, p 値が大きい (帰無仮説 H_0 は正しい) か小さい (帰無仮説 H_0 は間違い) かを決めるボーダーライン
 - ▶ **検出力 ($1-\beta$)** : 帰無仮説 H_0 が間違っているときに, 検定の結果, ちゃんと「帰無仮説 H_0 が間違っている」とする確率
- ▶ 帰無仮説 H_0 が正しいのに, 検定の結果「帰無仮説 H_0 は間違い」とするのは良くないと考え, 有意水準 α の値は小さい値 (例えば 5%) で固定し, その上で, 検出力 $1-\beta$ ができるべく大きくなるようにする



データ「DEP」の薬剤間の比較を行う場合・・・

1. 薬剤 C との比較のみに興味がある場合は、
「薬剤 A vs 薬剤 C」 「薬剤 B vs 薬剤 C」 の 2 種類の比較が出来る
2. 全ての比較結果について興味がある場合は、「薬剤 A vs 薬剤 B」
「薬剤 A vs 薬剤 C」 「薬剤 B vs 薬剤 C」 の 3 種類の比較が出来る

- ▶ α を 5% とする
- ▶ 1 回の検定を行う場合に Type I error を起こす確率は 5%
- ▶ 上記 2 の様に 3 回の検定を行う場合「少なくとも 1 回 Type I error を起こす確率」は、以下の計算から 14.3% となる
 - ▶ 少なくとも 1 回 Type I error を起こす確率
= $1 - (1 \text{ 回も Type I error を起こさない確率})$
= $1 - (1 - \alpha)^3 = 1 - (1 - 0.05)^3 = 0.1426 \quad 14.3\%$



検定回数と「少なくとも 1 回 Type I error を起こす確率」

- ▶ 確率が 14.3% だと、「検定」の手順 5「得られた確率が非常に小さい場合」に合致するかどうかがあやしくなる
- ▶ ここで、検定回数と「少なくとも 1 回 Type I error を起こす確率」の関係を表す表を紹介する（ $\alpha = 5\%$ の場合）

検定回数	少なくとも 1 回 Type I error を起こす確率
1	5.0 %
2	9.8 %
3	14.3 %
5	22.6 %
10	40.1 %
20	64.2 %
50	92.3 %
100	99.4 %



イントロのまとめ

- ▶ 「1回の検定の Type I error (α)」は 5% だが「一連（複数回）の検定の Type I error」が 5% よりも大きくなってしまふという問題は「多重性の問題」と呼ばれる
- ▶ 1つのデータに対して複数回検定を行う際に、計画している検定を全て行ったとしても Type I error が 5% 以下となるようにしたい場合、すなわち複数回の検定全体で「少なくとも 1 回 Type I error を起こす確率」を 5% に調節したい場合は、検定 1 回あたりの有意水準 α を 5% よりも小さくして検定を行う必要がある
- ▶ 以降では、「多重性の問題」を解消するための多重性の調整方法を紹介する　大きく分けて以下の 2 つの方法
 - ① 検定 1 回あたりの有意水準 α を調節する
 - ② 有意水準 α はそのままが良い（が、検定の順番等に制約をかける場合も）



本日のメニュー

1. イントロ

2. 帰無仮説族と FWE, 採択と保留

3. And の論理, Or の論理, ボンフェローニの方法

4. 固定順検定

5. ダネットの方法とチューキーの方法

6. その他の方法



薬剤が 3 種類以上ある場合は帰無仮説が複数存在する

- ▶ データ「DEP」について、薬剤 A, B, C のパラメータ（平均値など）をそれぞれ μ_A , μ_B , μ_C とする
- ▶ このとき、パラメータの比較について以下の 4 個の帰無仮説が存在する

- ▶ $H_{(A,B,C)} : \mu_A = \mu_B = \mu_C$ 包括的帰無仮説
(overall null hypothesis)
- ▶ $H_{(A,B)} : \mu_A = \mu_B$
- ▶ $H_{(A,C)} : \mu_A = \mu_C$
- ▶ $H_{(B,C)} : \mu_B = \mu_C$

部分帰無仮説
(subset null hypothesis)

- ▶ そのうち、推測の対象としている部分帰無仮説に絞る
 - ▶ 薬剤 C との比較のみに興味がある場合（Dunnett型）
 $F_D = \{ H_{(A,C)}, H_{(B,C)} \}$
 - ▶ 全ての比較結果について興味がある場合（Tukey型）
 $F_T = \{ H_{(A,B)}, H_{(A,C)}, H_{(B,C)} \}$



帰無仮説族と Type I FWE

- ▶ F_D や F_T を「帰無仮説族（帰無仮説ファミリー）」と呼ぶ
- ▶ 薬剤 C との比較のみに興味がある場合を考える
 - ▶ $F_D = \{H_{(A,C)}, H_{(B,C)}\}$, 1 回の検定の有意水準 α を 5% とする
- ▶ パラメータ (μ_A, μ_B, μ_C) の真の配置が . . .
 - ① $\mu_A = \mu_B = \mu_C$ の場合
 - 2 回の検定のうち $\{H_{(A,B)}, H_{(A,C)}\}$ のいずれかが誤って棄却される確率
= $1 - (1 \text{ 回も Type I error を起こさない確率})$
= $1 - (1 - \alpha)^2 = 1 - (1 - 0.05)^2 = 9.8\%$
 - ② $\mu_A \neq \mu_C, \mu_B = \mu_C$ の場合
 - $H_{(A,C)}$ は正しくないので、これらが誤って棄却されることはない
 - 2 回の検定のうち $\{H_{(A,B)}, H_{(A,C)}\}$ のいずれかが誤って棄却される確率
= $H_{(B,C)}$ が誤って棄却される確率 = 5% (なので多重性の考慮は不要)



帰無仮説族と Type I FWE

- ▶ 帰無仮説族（例えば F_0 ）についてパラメータの真の配置を 1 つに固定したときに
「正しい帰無仮説のうち少なくとも 1 つが誤って棄却される確率」を Type I FWE (Type I Familywise error rate) とよぶ
- ▶ Type I FWE はパラメータの真の配置に依存する
が、「パラメータの真の配置」は誰にも分からない・・・
- ▶ そこで、全ての「パラメータの真の配置」を想定して Type I FWE を計算し、最大の Type I FWE を 5% にする必要がある



「採択する」と「保留する」

- ▶ 「 $H_0: \mu_A = \mu_B$ 」に関する検定の結果，有意差が出なかった場合

- ① 帰無仮説は棄却されなかった

- ② 帰無仮説を採択する 「 $\mu_A = \mu_B$ 」と結論する

②を積極的に主張することは出来ない！

- ▶ 統計では「帰無仮説 H_0 が間違っていない」場合でも
「帰無仮説 H_0 が正しい」と積極的に言えないという話がある

有意差が出ない理由は様々（次頁以降で 3 つの例を紹介）

臨床的に意味のある差，Type I error (α)，検出力 ($1-\beta$) 等を考慮して例数設計をした上でデータを取ったにも関わらず，

「帰無仮説 H_0 が間違っていない」場合は，例数設計をせずにデータを取った場合に比べて「帰無仮説を採択する」ことが言いやすくなる



3つの例：①検定結果は例数に依存

- ▶ QOL の平均値の差 = 2, 各薬剤共通の標準偏差 = 2, 各薬剤は同じ例数, $\alpha=5\%$, 等分散を仮定した 2 標本 t 検定を行う
- ▶ 「各薬剤の例数」だけを変えて検定を行う

検定結果は例数に依存する

各薬剤の例数	p
2	0.4226
3	0.2879
4	0.2070
5	0.1525
6	0.1139
7	0.0859
8	0.0653
9	0.0499
10	0.0382

有意差なし

有意差あり



3つの例：②検定結果はばらつきに依存

- ▶ QOL の平均値の差 = 2, 各薬剤は同じ例数, $\alpha=5\%$, 等分散を仮定した 2 標本 t 検定を行う
- ▶ 「各薬剤共通の標準偏差」だけを変えて検定を行う

検定結果は「各群共通の標準偏差」の値に依存する

標準偏差	p
9	0.6253
8	0.5830
7	0.5309
6	0.4657
5	0.3829
4	0.2783
3	0.1534
2	0.0382
1	0.0003

有意差なし

有意差あり



3つの例：③検定結果は有意水準 α に依存

- ▶ 検定の結果, p 値は全て 0.031
- ▶ 「検定 1 回あたりの有意水準 α 」だけを変えて検定結果を解釈する
結果は「検定 1 回あたりの有意水準 α 」の値に依存する

p	α
0.031	0.050
0.031	0.045
0.031	0.040
0.031	0.035
0.031	0.030
0.031	0.025
0.031	0.020
0.031	0.015
0.031	0.010

有意差あり

有意差なし



「採択する」と「保留する」

- ▶ 例数が少ない場合は，臨床的な差があるにも関わらず「有意差なし」となる可能性がある
- ▶ ばらつきが大きい場合は，臨床的な差があるにも関わらず「有意差なし」となる可能性がある
- ▶ 多重比較の場合は，さらに「検定 1 回あたりの Type I error を 5% よりも小さくする」ため，検出力が通常の検定の場合よりも小さい（「正しくない帰無仮説」が棄却され損う確率が大きくなっている）
- ★ 「帰無仮説を採択する (accept)」は積極的にいいにくいが，
多重比較を行った場合はもっといいにくい点に留意する
多重比較の場合は，帰無仮説族の中の 1 つの帰無仮説について検定を行ったときに有意でなかった場合は「帰無仮説を保留する (retain)」という表現（「採択」よりも弱い表現）を使う場合がある



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. **And の論理, Or の論理, ボンフェローニの方法**
4. 固定順検定
5. ダネットの方法とテューキーの方法
6. その他の方法



多重比較の手順

- ▶ 事前に検定の対象となる「部分帰無仮説」を「帰無仮説族」の中に列挙し、解析に用いる多重比較の方法も事前に決めておく必要がある
- ▶ 多重比較の方法によっては「検定 1 回あたりの Type I error」を調整する必要が**無い**場合がある
 - すなわち、検定 1 回あたり $\alpha=5\%$ で検定しても良い場合がある
- ▶ 2 つの帰無仮説 H_{01} , H_{02} について
 - ▶ H_{01} の検定統計量 $T_1 : T_1$ c_1 なら H_{01} を棄却
 - ▶ H_{02} の検定統計量 $T_2 : T_2$ c_2 なら H_{02} を棄却
 - ▶ $\Pr(T_i \leq c_i | H_{0i}) = 0.05$ ($i = 1, 2$), T_1 と T_2 は独立とする



多重比較の手順〔Andの論理〕

- ① 2つの帰無仮説 H_{01} , H_{02} がともに成立しているとき,
 H_{01} と H_{02} を両方とも誤って棄却する確率

$$\begin{aligned}\text{Type I error} &= \Pr(T_1 < c_1 \text{ and } T_2 < c_2) = \Pr(T_1 < c_1) \times \Pr(T_2 < c_2) \\ &= 0.05 \times 0.05 < 0.05\end{aligned}$$

- ② 帰無仮説 H_{01} のみ成立しているとき,
 H_{01} と H_{02} を両方とも誤って棄却する確率

$$\begin{aligned}\text{Type I error} &= \Pr(T_1 < c_1 \text{ and } T_2 < c_2) = \Pr(T_1 < c_1) \times \Pr(T_2 < c_2) \\ &= 0.05 \times \Pr(T_2 < c_2) < 0.05\end{aligned}$$

- ▶ ①, ②ともに 0.05 以下 「検定 1 回あたりの Type I error」を調整する必要は無い ($\alpha=5\%$ でよい)
- ▶ 実は T_1 と T_2 が独立でない場合も調整する必要が無い



多重比較の手順〔Orの論理〕

- ① 2つの帰無仮説 H_{01} , H_{02} がともに成立しているとき,
 H_{01} と H_{02} のいずれかを棄却する確率

$$\begin{aligned}\text{Type I error} &= \Pr(T_1 > c_1 \text{ or } T_2 > c_2) \\ &= 1 - \Pr(T_1 < c_1 \text{ and } T_2 < c_2) \\ &= 1 - \{0.95 \times 0.95\} = 0.0975 > 0.05\end{aligned}$$

- ▶ 「検定1回あたりの Type I error」を調整する必要があるある！
(1回の検定を $\alpha=5\%$ とすると Type I FWE が 5% を超える)
- ▶ では、この場合はどうやって多重性を調整する？
「検定1回あたりの Type I error」を調整する一番単純な方法が
ボンフェローニの方法（ただし、検出力が最も低い方法）



ボンフェローニの不等式： $\Pr\left(\bigcup_{i=1}^3 E_i\right) \leq \sum_{i=1}^3 \Pr(E_i)$

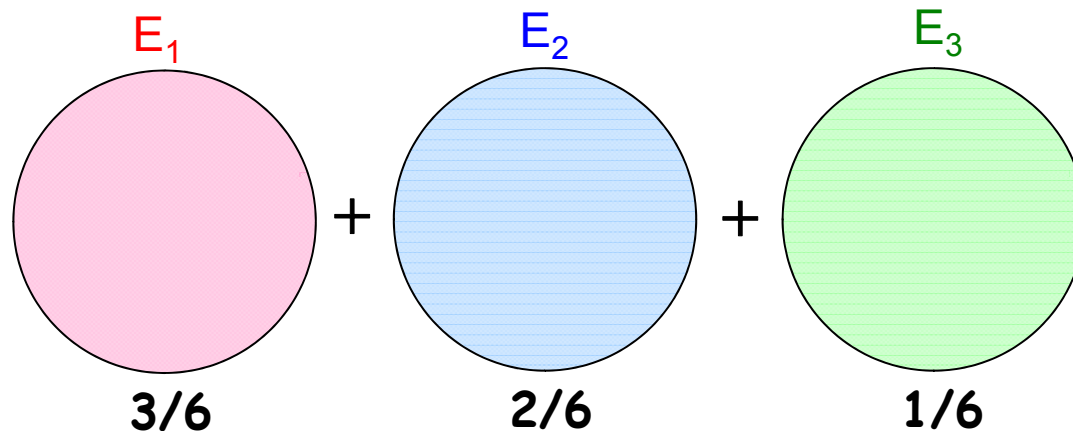
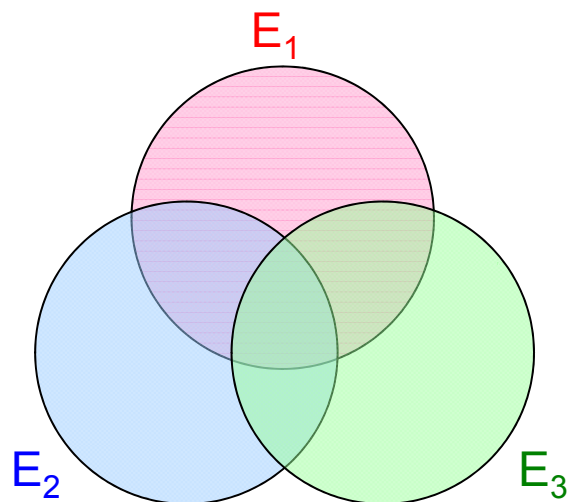
$$\Pr(E_1 + E_2 + E_3) \leq \Pr(E_1) + \Pr(E_2) + \Pr(E_3)$$

左辺：3個の事象 E_i のうち、少なくとも1つが成立する確率

右辺：各事象の成立確率の合計

例：サイコロを1回振ることを考える

$E_1 = 2$ の倍数, $E_2 = 3$ の倍数, $E_3 = 6$ の倍数



▶ 24 確率：4/6



ボンフェローニの方法

検定の順番を決めない場合（「Or」の場合）

- ▶ 1 回ごとの検定で求めた p 値を調整する
- ▶ 1 回ごとの検定における有意水準 α (5%) を調整する

ボンフェローニの方法

★ 検定を 3 回行う場合：

- ▶ Type I FWE $\alpha/3 + \alpha/3 + \alpha/3 = \alpha$ となるので
- ▶ 「検定 1 回あたりの Type I error」を $\alpha/3$ にする, 又は α をそのままにして p 値を 3 倍にすれば多重性の問題は解消出来る

★ 検定を m 回行う場合は 2 つの方法がある：

- ① α を 5% のままにして p 値を m 倍にする **今回はこれがメイン**
 - ② 検定 1 回あたりの Type I error を α/m にする
- ① 検定回数を k 回, i 回目の検定結果の p 値を p_i ($1 \leq i \leq k$) とすると
 $\hat{p}_i = \min(k p_i, 1)$ (1 を超えた場合は 1)



ボンフェローニの方法①

```
> pairwise.t.test(DEP$QOL, DEP$GROUP, p.adjust.method="none", pool.sd=F, var=T)
```

Pairwise comparisons using t tests with non-pooled SD

```
  A      B  
B 0.04728 -  
C 0.00057 0.14846
```

```
> p <- c(0.04728, 0.00057, 0.14846)
```

```
> p.adjust(p, method="bonferroni")
```

```
[1] 0.14184 0.00171 0.44538
```

薬剤 A の平均	薬剤 B の平均	薬剤 C の平均	2 標本 t 検定	ボンフェローニの方法による調整 p 値
6.5	4.0		4.7 %	14.1 %
6.5		2.5	< 0.01 %	0.1 %
	4.0	2.5	14.8 %	44.5 %



ボンフェローニの方法①

```
> pairwise.t.test(DEP$QOL, DEP$GROUP, p.adjust.method="bonferroni", pool.sd=F,  
var=T)
```

Pairwise comparisons using t tests with non-pooled SD

data: DEP\$QOL and DEP\$GROUP

```
  A      B  
B 0.1419 -  
C 0.0017 0.4454
```

P value adjustment method: bonferroni

薬剤 A の平均	薬剤 B の平均	薬剤 C の平均	2 標本 t 検定	ボンフェローニの方法による調整 p 値
6.5	4.0		4.7 %	14.1 %
6.5		2.5	< 0.01 %	0.1 %
	4.0	2.5	14.8 %	44.5 %



【参考】ボンフェローニの方法②

- ▶ 検定 1 回あたりの Type I error を α/m にすることも多重性の問題は解消できる
- ▶ 方法としては、とりあえず対象となる比較を（多重性の調整なしで）全て行い、「p 値が有意水準 $\alpha/3$ よりも小さいかどうか」で有意かどうかを判定する

結果は「ボンフェローニの方法①」と同じ

薬剤 A の平均	薬剤 B の平均	薬剤 C の平均	2 標本 + 検定	有意水準 ($\alpha/3$)
6.5	4.0		4.7 %	1.67 %
6.5		2.5	< 0.01 %	1.67 %
	4.0	2.5	14.8 %	1.67 %



【参考】 p.adjust.method に指定できる手法

p.adjust.method	手法
"bonferroni"	ボンフェローニの方法
"BH" , "BY"	FDR を調整する方法
"hochberg"	Hochberg の方法
"holm"	ホルムの方法
"hommel"	Hommel の方法
"none"	多重性の調整を行わない

▶ FDR (False Discovery Rate) :

例えば、遺伝子解析では帰無仮説が山ほど（数千～数万）あるので、
「最大の Type I FWE を 5% にする」と見逃しが多くなってしまふ

そこで、少し緩い基準として： $FDR = E(V/R)$ を 5% に制御する

V : (未知の)誤って棄却された帰無仮説の数, R : 棄却された帰無仮説の数



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. And の論理, Or の論理, ボンフェローニの方法
4. **固定順検定**
5. ダネットの方法とテューキーの方法
6. その他の方法



固定順検定

- ▶ QOL の平均値を薬剤間で複数回検定することを考える
- ▶ ここでは「薬剤 C との比較（2 種類の比較）のみに興味がある」とする
 - ▶ 薬剤 A vs 薬剤 C
 - ▶ 薬剤 B vs 薬剤 C
- ▶ ここで、検定の順序をあらかじめ以下のように決めておく
 1. 「薬剤 A vs 薬剤 C」の検定を行い、有意差があった場合のみ 2. に進む
（有意水準 $\alpha = 5\%$ ）
 2. 「薬剤 B vs 薬剤 C」の検定を行う（有意水準 $\alpha = 5\%$ ）
- ▶ 上記のように「有意差があった場合のみ次の検定を行う」という手順を「固定順検定（fixed sequence test）」という



固定順検定

- ▶ 1回の検定の有意水準 α を 5% にしているので、全体の有意水準 FWE (少なくとも 1 回 Type I error を起こす確率) が 5% を超えてしまう場合があるのでは? という懸念があるが . . .
 - ▶ 「1 番目の検定を行い、有意差があった場合のみ 2 番目の検定を行う」のがミソで、こうすることにより、1 回の検定の有意水準 α を 5% にしたとしても、全体の有意水準も 5% に抑えられることが知られている
 - ▶ 「有意差があった場合のみ 2 番目の検定を行う」ことをせずに、1 番目の検定結果によらず 2 つの検定を行った場合は、「Or の論理」により 1 回の検定の有意水準 α を調整 (5% よりも小さく) する必要がある
- ★ 固定順検定の欠点は、有意でなかった場合は以降の検定が出来ない (全て有意でなかったとなる) ので、検定の順番によっては損をする



固定順検定の例①

順番	薬剤 A の平均	薬剤 B の平均	薬剤 C の平均	p 値	解釈
1	6.5	4.0	2.5	4.7%	有意なので、次の検定を行う
2	6.5	4.0	2.5	< 0.1%	有意差がある

▶ 順番を以下のように決める

1. 薬剤 A vs 薬剤 C
2. 薬剤 B vs 薬剤 C

▶ 解釈の順番は以下の通り

1. 検定の結果, p 値は 4.7 % (有意水準 $\alpha = 5\%$) 有意なので 2. に進む
2. 検定の結果, p 値は 0.1 %未満 有意



固定順検定の例②

順番	薬剤 A の平均	薬剤 B の平均	薬剤 C の平均	p 値	解釈
1	6.5	<u>5.0</u>	2.5	14.1%	有意ではないので、ここで終了
2	6.5	<u>5.0</u>	2.5	< 0.1%	この検定結果は無効（見てはダメ） レポートする際は 14.1% とする

▶ 順番を以下のように決める

1. 薬剤 A vs 薬剤 C
2. 薬剤 B vs 薬剤 C

▶ 解釈の順番は以下の通り

1. 検定の結果, p 値は 14.1 % (有意水準 $\alpha = 5\%$) 有意ではないので終了
レポートする際は, 以降の検定結果は全て 14.1% とするのが作法の一つ
2. もし検定した場合は p 値は 0.1 %未満だが, この結果は見てはダメ



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. And の論理, Or の論理, ボンフェローニの方法
4. 固定順検定
5. **ダネットの方法とチューキーの方法**
6. その他の方法



ダネットの方法とチューキーの方法

1. ある薬剤（例：薬剤 C）とその他全部との比較に興味がある場合：
「薬剤 A vs 薬剤 C」「薬剤 B vs 薬剤 C」の 2 種類の比較が出来る
2. 全ての比較結果について興味がある場合：
「薬剤 A vs 薬剤 B」「薬剤 A vs 薬剤 C」「薬剤 B vs 薬剤 C」の
3 種類の比較が出来る

1. ダネットの方法を適用する（適切に多重性の調整が行われる）
2. チューキーの方法を適用する（適切に多重性の調整が行われる）

```
> library(multcomp) # パッケージの呼び出し
> DEP$GROUP <- relevel(DEP$GROUP, ref="C") # 「ある薬剤」のカテゴリ指定
> result <- lm(QOL ~ GROUP, data=DEP) # 回帰分析を実行
> result2 <- glht(result, linfct=mcp(GROUP="手法名")) # 多重性の調整
> summary(result2) # 結果の要約
> confint(result2) # 同時信頼区間
```



ダネットの方法

- ▶ ダネットの方法で多重性の調整を行った上で、QOL の平均値について「薬剤 A vs 薬剤 C」「薬剤 B vs 薬剤 C」の 2 種類の比較を行う

```
> library(multcomp) # パッケージの呼び出し
> DEP$GROUP <- relevel(DEP$GROUP, ref="C") # カテゴリのベースを「C」に変更
> result <- lm(QOL ~ GROUP, data=DEP)
> result2 <- glht(result, linfct=mcp(GROUP="Dunnett"))
> summary(result2)
```

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Dunnett Contrasts

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)	
A - C == 0	4.000	1.103	3.625	0.0012	**
B - C == 0	1.500	1.103	1.359	0.3011	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

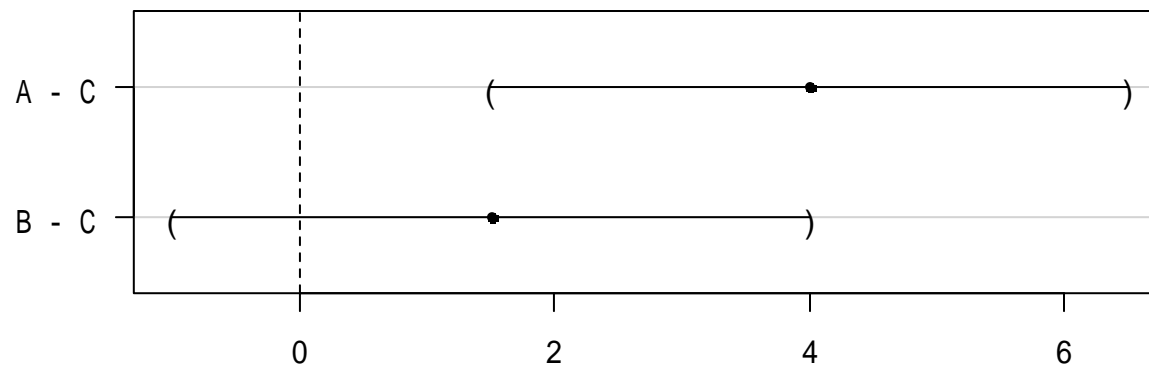


ダネットの方法

▶ 同時信頼区間の表示

```
> confint(result2)
95% family-wise confidence level
Linear Hypotheses:
      Estimate lwr      upr
A - C == 0  4.0000  1.4973  6.5027
B - C == 0  1.5000 -1.0027  4.0027
> plot(result2)
```

95% family-wise confidence level





テューキーの方法

- ▶ テューキーの方法の方法で多重性の調整を行った上で、QOL の平均値について「A vs B」「A vs C」「B vs C」の 3 種類の比較を行う

```
> result <- lm(QOL ~ GROUP, data=DEP)
> result2 <- glht(result, linfct=mcp(GROUP="Tukey"))
> summary(result2)
```

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: Tukey Contrasts

Linear Hypotheses:

	Estimate	Std. Error	t value	Pr(> t)	
A - C == 0	4.000	1.103	3.625	0.00175	**
B - C == 0	1.500	1.103	1.359	0.36891	
B - A == 0	-2.500	1.103	-2.266	0.06899	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

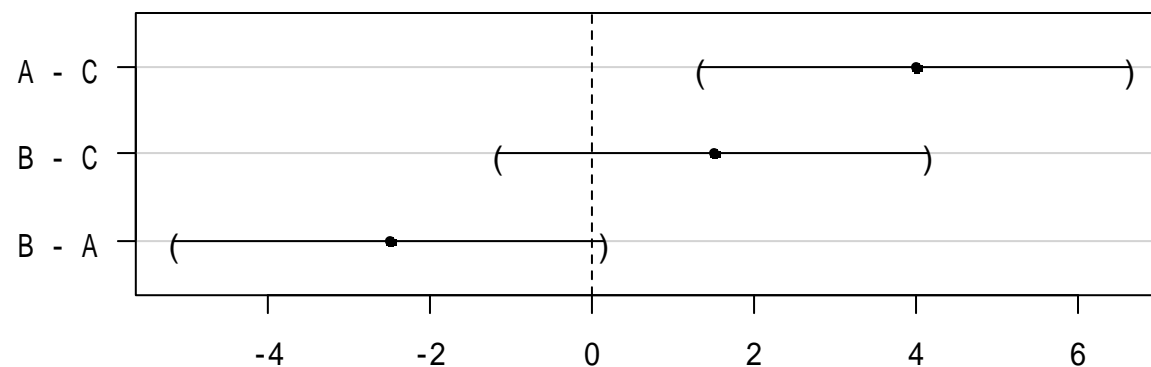


テューキーの方法

▶ 同時信頼区間の表示

```
> confint(result2)
95% family-wise confidence level
      Estimate lwr      upr
A - C == 0  4.0000  1.3447  6.6553
B - C == 0  1.5000 -1.1553  4.1553
B - A == 0 -2.5000 -5.1553  0.1553
> plot(result2)
```

95% family-wise confidence level





【参考】割合やイベントについても同様に解析可

```
> result <- glm(Y ~ GROUP, family=binomial, data=DEP)
> result2 <- glht(result, linfct=mcp(GROUP="Dunnett")) # 割合：ダネットの方法
> summary(result2)
```

Linear Hypotheses:

	Estimate	Std. Error	z value	Pr(> z)
A - C == 0	1.504e+00	6.892e-01	2.182	0.0533 .
B - C == 0	-4.054e-16	7.303e-01	0.000	1.0000

```
> result <- coxph(Surv(DAY,Y) ~ GROUP, data=DEP)
> result2 <- glht(result, linfct=mcp(GROUP="Tukey")) # イベント：チューキーの方法
> summary(result2)
```

Linear Hypotheses:

	Estimate	Std. Error	z value	Pr(> z)
A - C == 0	0.8287	0.5338	1.552	0.264
B - C == 0	-0.1699	0.6376	-0.266	0.961
B - A == 0	-0.9985	0.5355	-1.865	0.147



パッケージ「multcomp」に用意されている手法

p.adjust.method	手法
"bonferroni"	ボンフェローニの方法
"BH" , "BY"	FDR を調整する方法
"hochberg"	Hochberg の方法
"holm"	ホルムの方法
"hommel"	Hommel の方法
"none"	多重性の調整を行わない
"Shaffer"	Shaffer の方法
"Westfall"	Westfall の方法

- ▶ 他にも, "single-step" や "free" 等が指定できる
(詳細は「Multiple Comparisons Using R」を参照)



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. And の論理, Or の論理, ボンフェローニの方法
4. 固定順検定
5. ダネットの方法とテューキーの方法
6. その他の方法



ホルムの方法

```
> raw.p <- c(0.300, 0.010, 0.020, 0.025)           # 調整前のp値
> p.adjust(raw.p, method="holm")                   # ホルムの方法
[1] 0.30 0.04 0.06 0.06
> p.adjust(raw.p, method="hochberg")               # Hochbergの方法
[1] 0.30 0.04 0.05 0.05
> p.adjust(raw.p, method="hommel")                 # Hommelの方法
[1] 0.30000000 0.03333333 0.04000000 0.05000000
```

- ▶ ホルム (Holm) の方法 (有意水準 α , 対象とする比較は k 個)
 1. 対象とした比較について全て検定を行った後, p 値を小さい順に並べる
 $p^{(1)} \quad \dots \quad p^{(k)}$
 2. 以下の式から i 番目の調整 p 値を求める (便宜上, $p^{(0)} = 0$ とする)
$$p_i = \min[1, \max\{(k-i+1)p^{(i)}, p_{i-1}\}] \quad (i = 1, \dots, k)$$
 3. p_1 から昇順に有意かどうかを判定し, 有意でない時点で検定終了
- ★ 上記の例で「0.30, 0.04, 0.06, **0.05**」となっていない点に注意



Hochberg (ホッフベルグ) の方法

```
> raw.p <- c(0.300, 0.010, 0.020, 0.025)           # 調整前のp値
> p.adjust(raw.p, method="holm")                   # ホルムの方法
[1] 0.30 0.04 0.06 0.06
> p.adjust(raw.p, method="hochberg")               # Hochbergの方法
[1] 0.30 0.04 0.05 0.05
> p.adjust(raw.p, method="hommel")                 # Hommelの方法
[1] 0.30000000 0.03333333 0.04000000 0.05000000
```

▶ Hochberg の方法 (有意水準 α , 対象とする比較は k 個)

1. 対象とした比較について全て検定を行った後, p 値を大きい順に並べる

$$p^{(k)} \quad \dots \quad p^{(1)}$$

2. 以下の式から i 番目の調整 p 値を求める ($p_k = p^{(k)}$ とする)

$$p_i = \min[1, \min\{(k-i+1)p^{(i)}, p_{i+1}\}] \quad (i = k-1, \dots, 1)$$

3. p_k から降順に有意かどうかを判定し, 有意である時点で検定終了 (以降の仮説は全て有意であると判定する)



Williams の方法

- ▶ QOL の平均値を薬剤間で複数回検定することを考える
- ▶ 「薬剤 C との比較のみに興味があり」，かつ単調性が仮定出来る状況 ($\mu_C \quad \mu_B \quad \mu_A$) であるとする
- ▶ 推測の対象となる帰無仮説族を規定し有意水準 α を 5% とする
 $F_W = \{ H_{(C,B,A)}, H_{(C,B)} \}$ 対応する対立仮説を $H^A_{(C,B,A)}$ ， $H^A_{(C,B)}$ とする
 1. $H_{(C,B,A)} : \mu_C = \mu_B = \mu_A$ ， $H^A_{(C,B,A)} : \mu_C \quad \mu_B \quad \mu_A$
(少なくとも一つの " " が "<")
 2. $H_{(C,B)} : \mu_C = \mu_B$ ， $H^A_{(C,B)} : \mu_C < \mu_B$
- ▶ 1. に関する検定が有意でなければ，検定終了 (何もいえない)
- ▶ 1. に関する検定が有意であれば 2. の検定を行う
 - ▶ 有意である : $\mu_C < \mu_B < \mu_A$
 - ▶ 有意でない : $\mu_C < \mu_A$



Williams の方法

```
> n <- c(20,20,20) # Williamsの方法で用いる
> -contrMat(n, type="Williams") # 対比係数 (減少傾向)
      Multiple Comparisons of Means: Williams Contrasts
      1      2      3
C 1 1  0.0 -1.0
C 2 1 -0.5 -0.5

> DEP$GROUP <- factor(DEP$GROUP, levels=c("A","B","C")) # 薬剤の順番を指定
> result <- lm(QOL ~ GROUP, data=DEP)
> result2 <- glht(result, alternative="less", # Williamsの方法
+               linfct=mcp(GROUP="Williams"))
> summary(result2)
Multiple Comparisons of Means: Williams Contrasts
Linear Hypotheses:
      Estimate Std. Error t value Pr(<t)
C 1 >= 0  -4.0000     1.1034  -3.625 0.000509 ***
C 2 >= 0  -3.2500     0.9556  -3.401 0.001003 **
```



本日のメニュー

1. イントロ
2. 帰無仮説族と FWE, 採択と保留
3. **And** の論理, **Or** の論理, ボンフェローニの方法
4. 固定順検定
5. ダネットの方法とテューキーの方法
6. その他の方法



参考文献

- ▶ **Multiple Comparisons Using R** (Frank Bretz et. al., CRC press)
- ▶ 治験の統計解析 (Alex Dmitrienko 他著, 田崎 武信 監訳, 講談社)
- ▶ 統計的多重比較法の基礎 (永田 靖 他, サイエンティスト社)
- ▶ 統計学 (白旗 慎吾 著, ミネルヴァ書房)
- ▶ **The R Tips** 第2版 (オーム社)
- ▶ **R 流!** イメージで理解する統計処理入門 (カットシステム)

R で統計解析入門

終